TE

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/044,782	01/11/2002	Andreas Arning	DE920000057US1	8377

25259 7590 07/27/2005

IBM CORPORATION
3039 CORNWALLIS RD.
DEPT. T81 / B503, PO BOX 12195
REASEARCH TRIANGLE PARK, NC 27709

EXAMINER	
LY, ANH	

ART UNIT	PAPER NUMBER
2162	

DATE MAILED: 07/27/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/044,782	ARNING ET AL.	
	Examiner	Art Unit	
	Anh Ly	2162	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *0502/2005*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-13* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-13* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *11 January 2002* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

1. This Office Action is response to Applicant's Response filed on 05/02/2005.

2. Claims 1-13 are pending in this Application.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

5. Claims 1-13 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent No. 6,529,891 issued to Heckerman in view of US Patent No. 6,295,504 issued to Ye et al. (hereinafter Ye).

With respect to claim 1, Heckerman teaches determining a foreground frequency of a bucket within a first cluster (abstract, a number of states of a variable: foreground frequency of a variable which can take a certain number of states: such as age (a range of age of a group of person) and gender of a person (male and female) are numerical and categorical variables: figs. 2, 12, 14 and 29, col. 5, lines 40-67 and col. 6, lines 1-40);

determining a background frequency of the bucket with respect to all of the clusters (a value the hidden variable can take: col. 7, lines 20-52, col. 8, lines 1-8 and col. 13, lines 10-58); and

comparing the foreground and background frequencies (comparing the variable C, the number of states of C in the number of clusters of a population: col. 37, lines 10-67 and col. 38, lines 1-20).

Heckerman teaches determining the number of clusters (abstract) determining and storing the variable and a certain number of states of each variable, a value the variable can take or a bucket of a cluster, such as age (a range of age of a group of person) and gender of a person (male and female) are numerical and categorical variable, the foreground and background of a bucket of a variable C, the number of states of C in the number of clusters of a population. Also Heckerman teaches determining the comparison of the scores of a variable C by iterating the parameter over the possible choices of the number of states of C (col. 37, lines 10-35). Heckerman does not clearly teach determining a quality index based on the comparison.

However, Ye teaches optimal number of clusters and quality indexes and determining for each new measurement point (col. 13, lines 18-30 , col. 14, lines 15-22 and col. 16, lines 35-50), and the neighboring index values are used to determine a Kernel representative Index for each measurement points (abstract, col. 15, lines 55-65).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to combine the teachings of Heckerman with the teachings of Ye, wherein the number of clusters and each cluster having a set of buckets of each variable in the system provided therein (Heckerman's figs. 29, 30), would incorporate the use of determining a quality index of each cluster number associated with the determination for each new measurement point, in the same conventional manner as disclosed by Ye (abstract, col. 13, lines 18-30 and col. 14, lines 15-22). The motivation being for easing to group clustered data into clusters such that belonging to the same cluster have a high degree of similarity.

With respect to claim 2, Heckerman teaches wherein said comparing step further comprises subtracting the relative foreground and background frequencies (the number of states of C in the number of clusters of a population: col. 37, lines 10-67 and col. 38, lines 1-20).

With respect to claim 3, Heckerman teaches squaring the result of the comparison (col. 14, lines 20-30).

With respect to claim 4, Heckerman teaches determining an optimal number of clusters (col. 37, lines 8-55); and

comparing the optimal number of clusters to the actual number of clusters resulting from the clustering date processing operation (col. 39, lines 30-45, col. 41, lines 46-67 and col. 42, lines 1-8).

With respect to claim 5, Heckerman teaches wherein the optimal number of clusters is determined by a maximum number of buckets for a variable (figs. 31 & 32, col. 40, lines 10-50).

With respect to claim 6, Heckerman teaches wherein the optimal number of clusters is set to a threshold value in case the maximum number of buckets is greater than the threshold value (threshold value: col. 40, lines 10-65).

With respect to claim 7, Heckerman teaches determining a factor based on the optimal number of clusters and the actual number of clusters and multiplying the result of the comparison of the relative foreground and background frequencies with the factor (col. 13, lines 16-65 and col. 40, lines 10-65).

With respect to claim 8, Heckerman teaches determining a normalizing value being independent of any correlations between fields of the data on which the data processing operation is applied and normalizing the result of the comparison of the foreground and background frequencies by means of the normalizing value (col. 26, lines 22-67 and col. 27, lines 1-28).

With respect to claim 9, Heckerman teaches comparing the background an imaginary cluster having a foreground frequency of bucket equal to one, comparing the background frequencies of the buckets with an imaginary cluster having a foreground frequency of the bucket equal to zero, and summing the results of the corresponding

comparison values (col. 41, lines 25-60; also comparing the variable C, the number of states of C in the number of clusters of a population: col. 37, lines 10-67 and col. 38, lines 1-20).

With respect to claim 10, Heckerman teaches performing a number of data clustering operation (col. 37, lines 8-35); and

selecting the result of the data clustering (col. 21, lines 12-35 also col. 19, lines 40-67).

Heckerman teaches determining the number of clusters (abstract) determining and storing the variable and a certain number of states of each variable, a value the variable can take or a bucket of a cluster, such as age (a range of age of a group of person) and gender of a person (male and female) are numerical and categorical variable, the foreground and background of a bucket of a variable C, the number of states of C in the number of clusters of a population. Also Heckerman teaches determining the comparison of the scores of a variable C by iterating the parameter over the possible choices of the number of states of C (col. 37, lines 10-35). Heckerman does not clearly teach determining a quality index for each result of the data clustering operations.

However, Ye teaches optimal number of clusters and quality indexes and determining for each new measurement point (col. 13, lines 18-30 , col. 14, lines 15-22 and col. 16, lines 35-50), and the neighboring index values are used to determine a Kernel representative Index for each measurement points (abstract, col. 15, lines 55-65).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to combine the teachings of Heckerman with the teachings of Ye, wherein the number of clusters and each cluster having a set of buckets of each variable in the system provided therein (Heckerman's figs. 29, 30), would incorporate the use of determining a quality index of each cluster number associated with the determination for each new measurement point, in the same conventional manner as disclosed by Ye (abstract, col. 13, lines 18-30 and col. 14, lines 15-22). The motivation being for easing to group clustered data into clusters such that belonging to the same cluster have a high degree of similarity.

With respect to claim 11, Heckerman teaches a selecting an initial set of clusters (fig. 10, col. 16, lines 60-67 and col. 17, lines 1-32); and

performing a number of iterations (col. 37, lines 8-35).

Heckerman teaches determining the number of clusters (abstract) determining and storing the variable and a certain number of states of each variable, a value the variable can take or a bucket of a cluster, such as age (a range of age of a group of person) and gender of a person (male and female) are numerical and categorical variable, the foreground and background of a bucket of a variable C, the number of states of C in the number of clusters of a population. Also Heckerman teaches determining the comparison of the scores of a variable C by iterating the parameter over the possible choices of the number of states of C (col. 37, lines 10-35). Heckerman does not clearly teach determining a quality index for the clusters.

However, Ye teaches optimal number of clusters and quality indexes and determining for each new measurement point (col. 13, lines 18-30 , col. 14, lines 15-22 and col. 16, lines 35-50), and the neighboring index values are used to determine a Kernel representative Index for each measurement points (abstract, col. 15, lines 55-65).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to combine the teachings of Heckerman with the teachings of Ye, wherein the number of clusters and each cluster having a set of buckets of each variable in the system provided therein (Heckerman's figs. 29, 30), would incorporate the use of determining a quality index of each cluster number associated with the determination for each new measurement point, in the same conventional manner as disclosed by Ye (abstract, col. 13, lines 18-30 and col. 14, lines 15-22). The motivation being for easing to group clustered data into clusters such that belonging to the same cluster have a high degree of similarity.

With respect to claim 12, Heckerman teaches a method as discussed in claim 11.

Heckerman teaches determining the number of clusters (abstract) determining and storing the variable and a certain number of states of each variable, a value the variable can take or a bucket of a cluster, such as age (a range of age of a group of person) and gender of a person (male and female) are numerical and categorical variable, the foreground and background of a bucket of a variable C, the number of states of C in the number of clusters of a population. Also Heckerman teaches determining the comparison of the scores of a variable C by iterating the parameter over

the possible choices of the number of states of C (col. 37, lines 10-35). Heckerman does not clearly teach determining a quality index for the modified clusters.

However, Ye teaches optimal number of clusters and quality indexes and determining for each new measurement point (col. 13, lines 18-30 , col. 14, lines 15-22 and col. 16, lines 35-50), and the neighboring index values are used to determine a Kernel representative Index for each measurement points (abstract, col. 15, lines 55-65).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to combine the teachings of Heckerman with the teachings of Ye, wherein the number of clusters and each cluster having a set of buckets of each variable in the system provided therein (Heckerman's figs. 29, 30), would incorporate the use of determining a quality index of each cluster number associated with the determination for each new measurement point, in the same conventional manner as disclosed by Ye (abstract, col. 13, lines 18-30 and col. 14, lines 15-22). The motivation being for easing to group clustered data into clusters such that belonging to the same cluster have a high degree of similarity.

Claim 13 is essentially the same as claim 1 except that it is directed to a computer program product rather than a method (), and is rejected for the same reason as applied to the claim 1 hereinabove.

Contact Information

6. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Anh Ly whose telephone number is (571) 272-4039 or via E-Mail: ANH.LY@USPTO.GOV **or fax to (571) 273-4039**. The examiner can normally be reached on TUESDAY – THURSDAY from 8:30 AM – 3:30 PM.
If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Breene, can be reached on (571) 272-4107 or **Primary Examiner Jean Corrielus (571) 272-4032.**

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see **http://pair-direct.uspto.gov**. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). Any response to this action should be mailed to: Commissioner of Patents and Trademarks, Washington, D.C. 20231, or faxed to: Central Fax Center **(571) 273-8300**

JEAN M. CORRIELUS
PRIMARY

ANH LY
JUL. 22nd, 2005